# A key to **unsupervised semantic change modeling** is to **reduce cluster granularity**

## Similarity-Based Cluster Merging for Semantic Change Modeling


**Christopher Brückner** *bruckner@ufal.mff.cuni.cz*
**Leixin Zhang** *l.zhang-5@utwente.nl*
**Pavel Pecina** *pecina@ufal.mff.cuni.cz*

CHARLES UNIVERSITY
ÚFAL


### AXOLOTL'24 baseline

- Embed old senses and modern usage examples
- Cluster modern usage embeddings
- **Greedily assign old senses to clusters**

### Our approach

- Embed old senses and modern usage examples
- Cluster modern usage embeddings
- **Non-greedily assign old senses to clusters: Merge clusters if multiple good candidates exist**
- **Repeat the previous step for remaining clusters: Use cluster centroids as novel sense embeddings**

### Method

#### Pass 1: Word Sense Disambiguation

- Given old sense inventory $S$ and threshold $\tau$, for each cluster $C$ with centroid $\bar{\mathbf{c}}$, assign a cluster sense

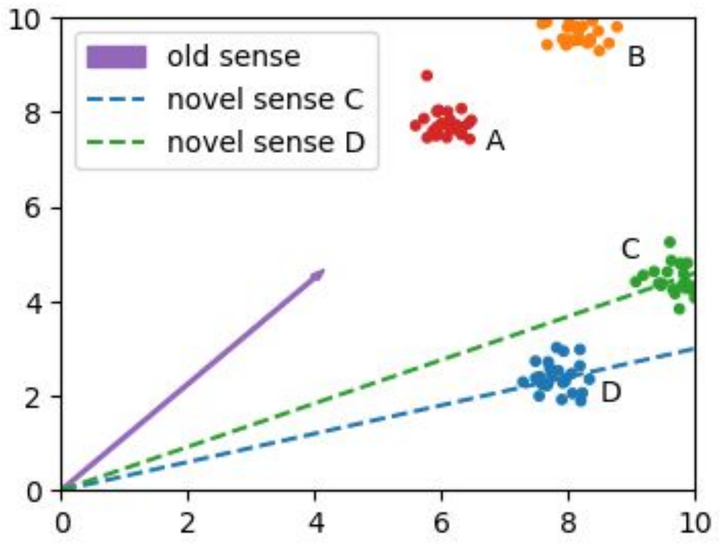
$$s_C = \begin{cases} argmax_{s \in S} \cos(\mathbf{s}, \bar{\mathbf{c}}), & \cos(\cdot) \geq \tau \\ s_{novel}, & otherwise \end{cases}$$

#### Pass 2: Word Sense Induction

- Induce and merge induced novel senses by updating the index of each remaining cluster $C_i$

$$i \leftarrow \begin{cases} argmax_{j > i} \cos(\bar{\mathbf{c}}_i, \bar{\mathbf{c}}_j), & \cos(\cdot) \geq \tau \\ i, & otherwise \end{cases}$$

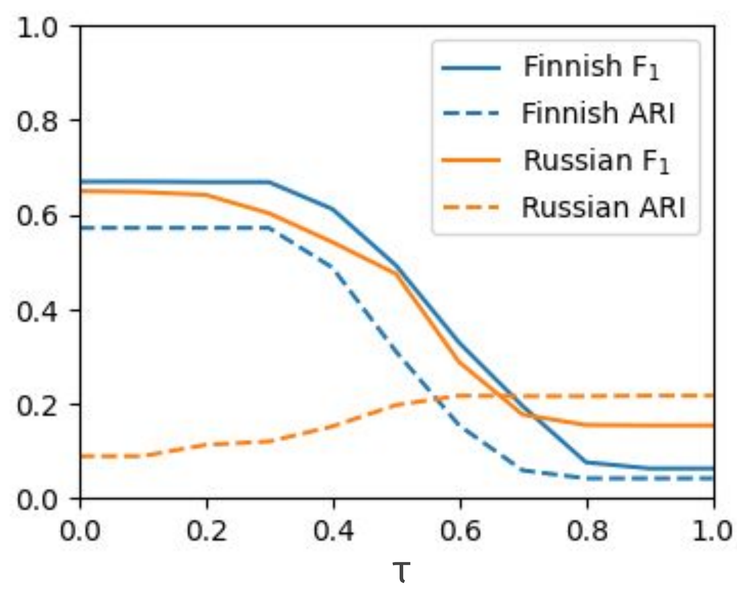
#### Example




Pass 1: Merge clusters A and B, assign the same old sense ID
Pass 2: Merge clusters C and D, assign the same novel sense ID

### Experiments

#### Scores for different τ




#### Results (avg. Fi-Ru-De)

| Team | ARI | F1 |
|---|---|---|
| *deep-change* | 0.413 | 0.750 |
| *Holotniekat (ours)* | 0.335 | 0.641 |
| *TartuNLP* | 0.310 | 0.590 |
| *IMS_Stuttgart* | 0.287 | 0.431 |
| *ABDN-NLP* | 0.221 | 0.487 |
| *WooperNLP* | 0.187 | 0.316 |
| *Baseline* | 0.041 | 0.207 |

#### Future work

- Use different thresholds to merge old and novel sense clusters
- Fuse the two passes instead of handling them as separate tasks


***Presented at LChange'24, online, Bangkok.***

*Supported by Charles University SVV project number 260 698 and Horizon Europe grant agreement number 101061016.*


***https://github.com/chbridges/axolotl24***